

Taylor & Francis
Taylor & Francis Group

Journal of Biomolecular Structure and Dynamics

ISSN: 0739-1102 (Print) 1538-0254 (Online) Journal homepage: https://www.tandfonline.com/loi/tbsd20

QSAR studies on the human sirtuin 2 inhibition by non-covalent 7,5,2-anilinobenzamide derivatives

Glaucio Monteiro Ferreira, Juliana Gallottini de Magalhães, Vinícius Gonçalves Maltarollo, Thales Kronenberger, Arasu Ganesan, Flávio da Silva Emery & Gustavo Henrique Goulart Trossini

To cite this article: Glaucio Monteiro Ferreira, Juliana Gallottini de Magalhães, Vinícius Gonçalves Maltarollo, Thales Kronenberger, Arasu Ganesan, Flávio da Silva Emery & Gustavo Henrique Goulart Trossini (2019): QSAR studies on the human sirtuin 2 inhibition by non-covalent 7,5,2-anilinobenzamide derivatives, Journal of Biomolecular Structure and Dynamics, DOI: 10.1080/07391102.2019.1574603

To link to this article: https://doi.org/10.1080/07391102.2019.1574603

Published online: 21 Feb 2019.

Submit your article to this journal

Article views: 16

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tbsd20

JOURNAL OF BIOMOLECULAR STRUCTURE AND DYNAMICS
https://doi.org/10.1080/07391102.2019.1574603

Taylor & Francis
Taylor & Francis Group

Check for updates

QSAR studies on the human sirtuin 2 inhibition by non-covalent 7,5,2-anilinobenzamide derivatives

Glaucio Monteiro Ferreira[a]*, Juliana Gallottini de Magalhães[a]*, Vinícius Gonçalves Maltarollo[b], Thales Kronenberger[c], Arasu Ganesan[d], Flávio da Silva Emery[e] and Gustavo Henrique Goulart Trossini[a]

[a]Department of Pharmacy, Faculty of Pharmaceutical Sciences, University of Sao Paulo (USP), Sao Paulo, SP, Brazil; [b]Department of Pharmaceutical Products, Faculty of Pharmacy, Federal University of Minas Gerais (UFMG), Belo Horizonte, Brazil; [c]Department of Internal Medicine VIII, University Hospital Tübingen, Tübingen, Germany; [d]School of Pharmacy, University of East Anglia, Norwich, UK; [e]Department of Pharmaceutical Sciences, Faculty of Pharmaceutical Sciences of Ribeirão Preto, University of São Paulo, Ribeirão Preto, SP, Brazil

Communicated by Ramaswamy H. Sarma

ABSTRACT

Sirtuin 2 is a key enzyme in gene expression regulation that is often associated with tumor proliferation control and therefore is a relevant anticancer drug target. Anilinobenzamide derivatives have been discussed as selective sirtuin 2 inhibitors and can be developed further. In the present study, hologram and three-dimensional quantitative structure–activity relationship (HQSAR and 3D-QSAR) analyses were employed for determining structural contributions of a compound series containing human sirtuin-2-selective inhibitors that were then correlated with structural data from the literature. The final QSAR models were robust and predictive according to statistical validation (q^2 and r^2_{pred} values higher than 0.85 and 0.75, respectively) and could be employed further to generate fragment contribution and contour maps. 3D-QSAR models together with information about the chemical properties of sirtuin 2 inhibitors can be useful for designing novel bioactive ligands.

Abbreviations: ANOVA: analysis of variance; CoMFA: comparative molecular fields analysis; CoMSIA: comparative molecular similarity indices analysis; cpd: compound; HL: hologram length; HQSAR: Hologram quantitative structure–activity relationship; MAE: mean absolute error; QSAR: quantitative structure–activity relationship; SIRT2: sirtuin 2; ρc: concordance correlation coefficient

ARTICLE HISTORY
Received 26 June 2018
Accepted 16 January 2019

KEYWORDS
Anilinobenzamide; CoMFA; CoMSIA; epigenetic; HQSAR; sirtuin

Introduction

Epigenetic covalent modifications of DNA and chromatin proteins strongly affect gene expression and cellular activity. Misregulation of these processes is relevant in several diseases where regulation of cell cycle and gene expression is fundamental, such as cancer, obesity and Parkinson's disease (Choudhary, Weinert, Nishida, Verdin & Mann, 2014; Feinberg & Vogelstein, 1983; Menzies, Zhang, Katsyuba & Auwerx, 2016; Rodriquez et al., 2006). A process such as acetylation of histone's lysine residues can neutralize their positive charge, which is responsible for interaction with DNA, leading to relaxation of the chromatin. Ultimately, regulating the availability of DNA to bind transcription proteins modulates the transcription level in that particular region of the genome (Menzies et al., 2016).

In this context, the enzymes responsible for the reversible process of acetylation of lysine residues in DNA binding proteins such as histones become a promising molecular target for drug discovery. Namely, histone acetyltransferases and histone deacetylases (HDACs) comprise a large family of enzymes that potentially can be exploited for drug discovery. The HDACs can be divided into two subgroups: zinc-dependent enzymes (HDAC1–11); and NAD^+-dependent enzymes, the so-called silent information regulators or sirtuins (SIRT1–7) (Choudhary et al., 2014).

Each sirtuin harbours a conserved catalytic central domain that is approximatey 270 amino acids long. Sirtuins catalyse cleavage of the glycosidic bond between nicotinamide (Rodriquez et al., 2006) and ADP-ribose of NAD^+, followed by transfer of the acetyl residue from the former acetyl-lysine-peptide to ADP-ribose, generating nicotinamide, deacetylated peptide and a mixture of 2′- and 3′-O-acetyl-ADP-ribose. The mechanism of this acetyl cleavage by sirtuins is as follows (see Figure 1 for an overview of the activity): (i) attack of acetyl-lysine at the ribose carbon bonded to nicotinamide, followed by neutralization of its positive charge after displacement of the nicotinamide in an SN_2-like reaction, which yields an O-alkylamidate adduct; (ii) the electron pair of 2′-hydroxyl reacts in an intramolecular nucleophilic attack on the carbon attached to the imine N atom, which leads to the formation of an acetal species; and (iii) hydrolysis of the

CONTACT Gustavo Henrique Goulart Trossini trossini@usp.br Department of Pharmacy, Faculty of Pharmaceutical Sciences, University of Sao Paulo (USP), Av. Lineu Prestes 580, Sao Paulo, SP 05508-900, Brazil
*These authors contributed equally to this work.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/07391102.2019.1574603

© 2019 Informa UK Limited, trading as Taylor & Francis Group

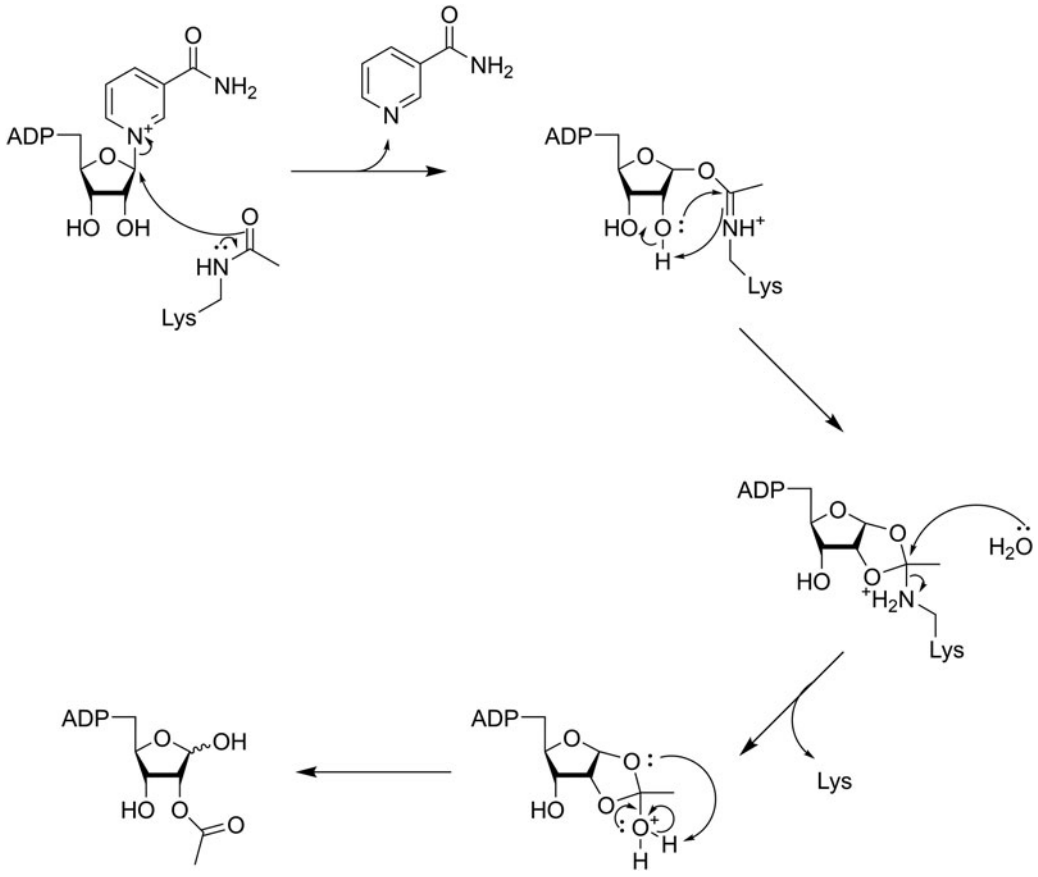


Figure 1. Sirtuin 2 (SIRT2) overview of the mechanism of deacetylation involved in the epigenetic regulation focused on this work. *N*-ε-acetylated proteins are deacetylated consuming one NAD^+ molecule during the process and releasing one free acetyl group.

acetal group releases the previously formed adduct, which is followed by an intramolecular transesterification that results in hydroxyacetyl and hydroxyl groups.

The seven human sirtuin enzymes have distinct cellular localizations and targets. SIRT1,6,7 are predominantly nuclear and active on histones and transcription factors, whereas SIRT3–5 have been described as mitochondrial (van de Ven, Santos & Haigis, 2017). As an example of the overall effect that SIRT modulation may have, we highlight SIRT1 activation, which promotes catabolism by deacetylating key mitochondrial target genes, such as peroxisome proliferator-activated receptor gamma (PPAR-γ) coactivator 1-alpha (PGC-1α) (Menzies et al., 2016; Rodgers et al., 2005). PGC-1α is a coactivator protein of PPAR-γ, the master regulator of metabolism in brown fat (Puigserver et al., 1998) and liver, with a fundamental role in cases of diabetes mellitus (Yoon et al., 2001). On the other hand, SIRT2 is unique because, despite nuclear localization, it can be found mainly in the cytosol (Ramis, Esteban, Miralles, Tan & Reiter, 2015), colocalizing with the microtubule network. SIRT2 targets FOXO1, suppressing adipocyte differentiation (Bobrowska, Donmez, Weiss, Guarente & Bates, 2012; Jiang et al., 2011; Jing, Gesta & Kahn, 2007). ATP-citrate lyase (ACL) inhibits lipid synthesis (Lin et al., 2013) and the cytosolic phosphoenolpyruvate carboxykinase (PEPCK-C^{mus}) stimulates hepatic gluconeogenesis (Jiang et al., 2011).

SIRT2 has had quite a controversial role, being both a tumour suppressor (Kim et al., 2011; Serrano et al., 2013) and also associated with tumorigenesis (Yang et al., 2013), but recent findings have proposed that its role in metastasis is related to the increase in cell motility of cancer cells (Mellini et al., 2017). With regard to importance as a regulator of metabolism and potential in cancer treatment, SIRT inhibitors have been described with high potency but low selectivity (Rotili et al., 2012). Inhibitors such as salermide (Lara et al., 2009) and EX-527 (Napper et al., 2005) are potent but lack SIRT2 selectivity, whereas others such as AGK2 are selective but comparatively less potent. SIRT inhibition is often associated with nicotinamides and benzamides, the mechanism of action being based on competition between substrates for binding into NAD^+ or nicotinamide pockets (Mai et al., 2005; Mellini et al., 2017; Napper et al., 2005; Suzuki et al., 2012).

Molecular modelling studies in terms of QSAR are relevant to guide the early stages of drug discovery; this has been shown on several occasions when QSAR models successfully enriched libraries with inhibitory activity or selected molecules against cancer-related proteins (Alam & Khan, 2017; Braga et al., 2014; Chen et al., 2017; Zhao et al., 2017). Specifically, computational techniques for SIRT modulation are relevant to guide the early stages of computer-aided drug discovery of selective new drug candidates (Andreoli & Del Rio, 2015; Chuang, Chang, Lin & Yang, 2015). In this sense, with this work we studied a series of 3′-phenethyloxy-

2-anilinobenzamide derivatives as SIRT2 inhibitors by generating a diverse set of QSAR models with different approaches (Suzuki et al., 2009, 2012).

Material and methods

Compound set and training/test separation

Seventy-five inhibitors for human SIRT2, based on 2-anilinobenzamide derivatives, were selected from the literature to generate QSAR models (Table S1 in Supplementary Material) (Suzuki et al., 2012). Biological activity data were converted into pIC_{50} (–logIC_{50}) values, which is the log of the IC_{50} of SIRT2 inhibition represented as molar values.

Compounds were grouped according to molecular structural diversity, drug-like properties and range of biological activity. Molecular structural diversity and drug-like properties were clustered using hierarchical cluster analysis (HCA) with the complete linkage method and Euclidean distance implemented in Pirouette 3.1 (Infometrix, 1985; Kowalski, 1977). Compounds were separated into different groups according to the biological activity range, divided into intervals of one log unit. Molecular structure information was encoded from PubChem fingerprints, whereas drug-like properties were represented by the following descriptors: ClogP, number of H-bond acceptors (HBA) and donors (HBD), topological polar surface area (TPSA), number of rotatable bonds (nRotB) and molecular weight (MW). Fingerprint and descriptor calculations were performed using PaDel descriptors (Yap, 2011) and all data were normalized before HCA.

Dataset compounds were separated into training and test sets (4:1). SIRT2 inhibition models were made with 75 compounds (60 and 15 compounds in training and test sets, respectively). Compounds were randomly separated between training and test sets and the test set was accepted if the final selection remained representative from each cluster without compromising the proportion in the total set (Figure S1 in Supplementary Material).

Two-dimensional (2D) structures of the compounds were drawn using Marvin Sketch version 15.8.31 (ChemAxon, Budapest, Hungary). Three-dimensional (3D) structures were drawn in Maestro2017 version 4 (Madhavi Sastry, Adzhigirey, Day, Annabhimoju & Sherman, 2013), treated in terms of protonation and charge, and flexibly aligned against the co-crystallized conformation of compound 71 (Protein Data Bank ID: 5Y5N). Final conformers were adjusted in terms of charge (PM1-BCC semi-empirical method: Halgren, 1996) using the module NDDO implemented in Maestro.

Models were generated using three independent QSAR approaches to provide insight into SIRT2 inhibition: HQSAR, comparative molecular fields analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA). All the models were generated using molecules from the training set as input for partial least-squares (PLS) regression methods implemented in SYBYL-X 2.1.1 (Tripos, Inc., 2011). For each approach, several different models were inferred by varying the standard parameter settings.

Generation of HQSAR models

HQSAR is a two-dimensional method that correlates biological activity with the structural features of a set of compounds using molecular fingerprints of fragments, so-called holograms. Molecular fragments that generate holograms can have different lengths and be distinct in terms of fragment counting: atomic type (A), bonds (B), connections (C), hydrogen atoms (H), chirality (Ch) and/or hydrogen bond donors/acceptors (DA). Initially, HQSAR models were generated with all the possible fragment distinction combinations by considering atoms at least and, for each combination, using all standard options of hologram length (53, 59, 61, 71, 83, 97, 151, 199, 257, 307, 353 and 401 bins). For these models, the default fragment size (from 4 to 7 atoms) was employed. By varying the fragment size, further models were generated from the best models of the previous step: leave-one-out cross-validation regression coefficient (Q^2_{LOO}) >0.6. Fragment size variation was performed from 1 to 11 atoms in intervals of two, three and four atoms.

Generation of CoMFA and CoMSIA models

3D-QSAR models were generated for both CoMFA, using positively charged sp^3 carbon as the probe atom to calculate molecular interaction fields, and CoMSIA, using positively charged sp^3 hybridized carbon as the probe atom to calculate a range of different similarity indices. These models were generated using the molecular alignment of training set molecules with an initial grid spacing of 2 Å and an energy cut-off of 30 kcal/mol. For CoMFA models, molecular interaction fields were calculated with the probe atom and calculating steric and electrostatic interactions with ligands using Lennard–Jones and Coulomb energy terms. Specifically, for CoMSIA different combinations of similarity indices (steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor) were calculated and the combination that generated the best model in terms of the highest Q^2_{LOO} among its pairs was further optimized using focus. The best obtained CoMFA and CoMSIA models were focused by changing both the grid spacing (1 to 4 Å in steps of 0.5 multiplied by the original distance) and the weight factor (0.3 to 1.5 in steps of 0.2 multiplied by the standard deviation (SD) of the original model). Different combinations of weight factor and grid spacing were employed and the respective models generated were ranked by the Q^2_{LOO} values. The maximum numbers of principal components (PCs) used in QSAR models were based on the size of the dataset (75 compounds). For each individual model, the lowest number of PCs sufficient for explaining the system's variability was used in order to avoid overtraining, according to Occam's razor principle (Gaudio & Zandonade, 2001).

Validation of QSAR models

The best models from each approach were then selected according to their robustness, represented by Q^2_{LOO}. The five best models of each approach, ranked in terms of robustness

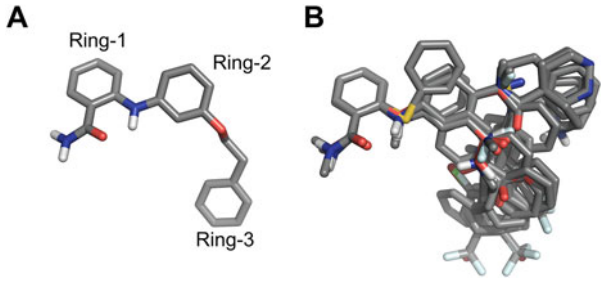


Figure 2. Overview of the data set. (A) Ligand from the co-crystallized SIRT2 (PDB ID: 5Y5N) used as a template for the alignment, where the three characteristic rings and the free-amid can be observed. (B) Final alignment of the molecules.

($Q^2_{LOO} > 0.6$), underwent external validation using their respective test set. Finally, the best models in terms of $Q^2_{(F2)}$ and $Q^2_{(F3)}$ (>0.7) were used to generate contribution and contour maps for the most and least active and selective compounds. Additional metrics, such as r_m^2, which compares correlation coefficients in the prediction of test compounds when passing through the origin (r_0^2) or intercepting any other point (r^2), were evaluated to assess the model predictability (a comprehensive review and detailed description of these metrics are provided by the works of Gramática and Roy (Chirico & Gramatica, 2011; Kunal et al., 2015). Additionally, the mean absolute error (MAE) was employed according to Roy et al. (Roy, Das, Ambure, & Aher, 2016) due to the limitations of R^2-based metrics, which is often dependent on the distribution of poorly diverse datasets. In this study, we followed the principles established by the OECD (Organisation for Economic Co-operation and Development) to develop reliable and predictive QSAR models: a defined endpoint, an unambiguous algorithm, defined applicability domain, appropriate validation and a mechanistic interpretation (Gramatica, 2007). The sensitivity of the models to chance correlations was checked by 80 runs of progressive scrambling (CoMFA and CoMSIA) and *y*-scrambling (HQSAR).

Molecular docking

Representative compounds were selected and underwent molecular docking on the *Hs*SIRT2 (Protein Data Bank ID: 5Y5N) active site to discuss possible mechanisms of interaction. The selected compound was prepared by adding hydrogen atoms (optimized by hydrogen bonding) and correcting missing side-chains, using the Protein Preparation Wizard followed by hydrogen bond assessment, both from Maestro2017 version 4 (Madhavi Sastry et al., 2013), with standard options. Only the protein–ligand complex hydrogen atoms were minimized. Water molecules with at least three hydrogen bonds with ligand and protein atoms, namely A1112, A1119 and A1129, were kept for docking.

Ligand structures for docking were prepared as mentioned previously for the alignment procedure. Docking calculations were performed using GOLD 5.2 (Jones, Willett, Glen, Leach & Taylor, 1997; Siligardi, Hussain, Patching & Phillips-Jones, 2014; Verdonk, Cole, Hartshorn, Murray & Taylor, 2003; Verdonk et al., 2005) within a cavity of 10 Å around the co-crystallized ligand and using ChemScore as the score function. Other scoring functions were evaluated for redocking purposes, resulting in poor performance (data not shown). Docking poses within the binding site were visualized and figures were generated using PyMOL (Delano, 2004).

Table 1. Statistical data of the best HQSAR models for SIRT2 inhibition.

Distinction	Fragment size	q^2	r^2	SEE	HL	*N*
A/C	5–8	0.90	0.95	0.16	257	5
A/C	6–9	0.90	0.95	0.17	257	4
A/C/Ch	6–9	0.89	0.95	0.17	257	4

Results and discussion

Data-set and alignment

Molecular alignment is a major step in MIF-based 3D-QSAR modelling due to the tri-dimensional nature of the descriptors, which are calculated based on a lattice grid (Wilkes, Stoyanova-Slavova, & Buzatu, 2016). In this case, alignment of the human SIRT2 inhibitors shows three important regions for this series that coincide with three rings (Figure 2A). Ring 1 is uniquely decorated with a conserved system of one H-bond acceptor and one H-bond donor (typically an amide group), whereas the two other rings have mostly hydrophobic substituents (Figure 2B). The hinge regions connecting these rings also have a conserved H-bond donor (between rings 1 and 2) and an H-bond acceptor (between rings 2 and 3).

HQSAR model generation

Among all the generated HQSAR models (Table S2 in Supplementary Material), the best SIRT2 inhibition models presented values of Q^2_{LOO} between 0.82 and 0.88. The model that best encompassed information comprising atoms and connections (A/C, with Q^2_{LOO} = 0.88), along with a second model adding the influence of chirality (A/C/Ch, Q^2_{LOO} = 0.88), underwent a cycle of further optimization, varying the fragment size, to generate models with improved statistical parameters (Table 1). The models had improved robustness after the parameter optimization step, which is in accordance with previous HQSAR studies (Seel, Turner & Willett, 1999). The best-optimized model, based on the Q^2_{LOO} values, was chosen for further interpretation. Also, models based on A/C were preferred instead of A/C/Ch, according to Occam's razor principle (Gaudio & Zandonade, 2001).

CoMFA and CoMSIA model generation

The best CoMFA models among all those calculated (Table S3 in Supplementary Material) had the same weight factor (1.0) with different grid distances (Table 2), all with Q^2_{LOO} values comparably higher than 0.80 and low error rates after focusing. The field contributions corroborated the similarity among the models. We could also observe that steric fields are more important than electrostatic fields for describing the structure–activity relationship for this set of molecules.

Table 2. Statistical data of the best constructed CoMFA models for SIRT2 inhibition and selectivity.

Focus							Field contribution	
d	W	q^2	SEV	r^2	SEE	N	S	E
No focus		0.84	0.31	0.99	0.03	10	0.79	0.21
0.7	1.0	0.90	0.26	0.98	0.09	15	0.77	0.23
1.0	1.0	0.90	0.25	0.99	0.08	15	0.77	0.23
1.5	1.0	0.90	0.25	0.93	0.09	15	0.77	0.23

d, distance factor; *w*, standard deviation weight factor; q^2, LOO cross-validation correlation coefficient; SEV, standard error of validation; *N*, optimal number of components; r^2, non-cross-validation correlation coefficient; SEE, standard error of estimation. Field contribution: *S*, steric; *E*, electrostatic.

Table 3. Statistical data of the best constructed CoMSIA models for SIRT2 inhibition and selectivity.

	Focus							Field contribution			
Fields	d	W	q^2	SEV	r^2	SEE	N	S	D	H	A
DAH	No focus		0.83	0.31	0.99	0.11	3	–	0.60	0.13	0.26
DAH	0.3	1.0	0.82	0.31	0.97	0.14	3	–	0.59	0.14	0.28
DAHS	No focus		0.81	0.32	0.98	0.10	2	0.23	0.41	0.11	0.23

d, distance factor; *w*, standard deviation weight factor; q^2, LOO cross-validation correlation coefficient; SEV, standard error of validation; *N*, optimal number of components; r^2, non-cross-validation correlation coefficient; SEE, standard error of estimation. Field contribution: *S*, steric; *E*, electrostatic.

Interestingly, the best CoMSIA model generated encompassed the information on hydrogen acceptors and donors, as well as hydrophobicity (Q^2_{LOO} = 0.83, Table 3 and Table S4 in Supplementary Material for all CoMSIA models), which complements the CoMFA models in terms of chemical information. In contrast to the CoMFA models, the focus procedure did not increase the model's robustness.

External validation and model selection

External validation, employing compounds from the test set, was made for the three best models of each method (Table 4). All models were considered satisfactory, with values within the specifications and according to OECD guidelines (Chirico & Gramatica, 2011; Gramatica, 2007; Satoh, Vudikaria, Kurano & Miyachi, 2005). All these three models are well predicted, as can be noticed by the r^2_{pred} values of >0.7 (Figure 3A and B, as a graphical representation) and the low root-mean-square error of prediction (RMSEP) rates. The high $Q^2_{(F2)}$ and $Q^2_{(F3)}$ values suggest the great predictability of SIRT2 inhibition. Additionally, the good proximity between predicted and observed activity for all selected models can be demonstrated by the parameter r^2_m, which also presented values of >0.75. Average r^2_m (>0.5) and Δr^2_m (<0.2) reinforce the good predictive potential of the previous metric. Moreover, the six chosen models can be considered reliable due to the concordance correlation coefficients (CCC) approaching unity (0.82–0.94). Residuals were always <1 and showed no correlation with predicted values (Figure S3 in Supplementary Material).

In order to test the robustness of the three selected models and stability for variation of training set composition, we also performed a leave-N-out (LNO) validation, with cross-validation group numbers varying from 5 to 50. The analyses were performed in triplicate and the final q^2 was calculated as the average q^2 obtained in all the runs. All the models showed average q^2 values of > 0.8 and low SD values, indicating good internal consistency (Table S5 in Supplementary Material). Among the validated models, one for each technique was chosen for visual interpretation.

Physicochemical interpretation of models

The HQSAR contribution maps (Figure 4) suggest that the 2-anilinobenzamide group has a positive contribution to SIRT2 inhibition. The HQSAR model also shows a negative contribution in the para position of ring 3 (as illustrated by compound 18) that is not observed in other compounds with larger linker regions (such as compound 71). It is worth mentioning that compound 18 has an extra carbon atom at the ring 2–3 linker that could restrain the conformation of ring 3 and therefore any *para* substituent.

Specifically, bulky substituents (in general a linker at the meta position and ring 3) extending from ring 2 could increase the SIRT2 inhibitory activity, but further ramifications on ring 3 are suggested not to be tolerated (Chuang et al., 2015), which is consistent with the hydrophobic contour field from the CoMSIA models (Figure 5). Accordingly, from CoMFA contour maps it is possible to note that ring 3 of compound 71 is located at a favorable steric region, in contrast to ring 3 of compound 18 or to compound 9 that has no ring (Figure 5). This difference could be attributed to the $-O(CH_2)_2-$ linker of the ring reaching the sterically favored region in compound 71, which allows the conformation of ring 3 to remain at a favorable position. This steric-disfavored contour would superimpose with Phe119, as shown in molecular dynamics simulations (Chuang et al., 2015), and corroborates the work of Suzuki: that methylene linker length is important to ensure inhibition of SIRT2 (Feldman, Baeza, & Denu, 2013; He, Hu, Zhang, & Lin, 2014).

Additionally, the presence of an oxygen atom between rings 2 and 3 was observed in active compounds and superimposes with favorable hydrophilic and donor fields from the CoMSIA models and with the favorable electrostatic contour from CoMFA models (Figure 5). Our CoMSIA contour maps point to the importance of having H-bond donor and acceptor groups near the linker region, external to ring 2.

With respect to the contour maps around compounds 71, 18 and 9, the green contours represent regions where bulky groups increase biological activity whereas yellow contours indicate areas where bulky groups decrease biological activity. Favorable electrostatic contributions are represented in blue and unfavorable contributions to biological activity are represented in red. Favorable acceptor contributions are highlighted in pink and unfavorable in grey. Similarly, favorable donor contributions are highlighted in purple with unfavorable in grey, and favorable hydrophobic contributions in orange with unfavorable in cyan. All favorable contours were set to level 80 and their unfavorable counterparts were set to level 20.

The substitutions in the *para* position of ring 3 observed in compounds **63**–**78** (Table S1 in Supplementary Material

Table 4. Validation of the three best models constructed from each method: HQSAR, CoMFA and CoMSIA.

Model	q^2	r^2_{pred}	RMSEP	r^2_m	$r^2_m{}'$	$\bar{r}^2_m$	Δr^2_m	$Q^2_{(F2)}$	$Q^2_{(F3)}$	CCC
HQSAR A/C (5–8)	0.90	0.79	0.32	0.73	0.70	0.72	0.03	0.79	0.80	0.90
HQSAR A/C (6–9)	0.90	0.87	0.25	0.86	0.81	0.84	0.05	0.87	0.88	0.94
HQSAR A/C/Ch (6–9)	0.89	0.87	0.25	0.86	0.80	0.83	0.07	0.87	0.88	0.94
CoMFA d 1.0 w 1.0	0.90	0.99	0.23	0.88	0.78	0.83	0.09	0.88	0.89	0.94
CoMSIA DAH No focus	0.83	0.99	0.28	0.63	0.79	0.71	0.15	0.83	0.84	0.90

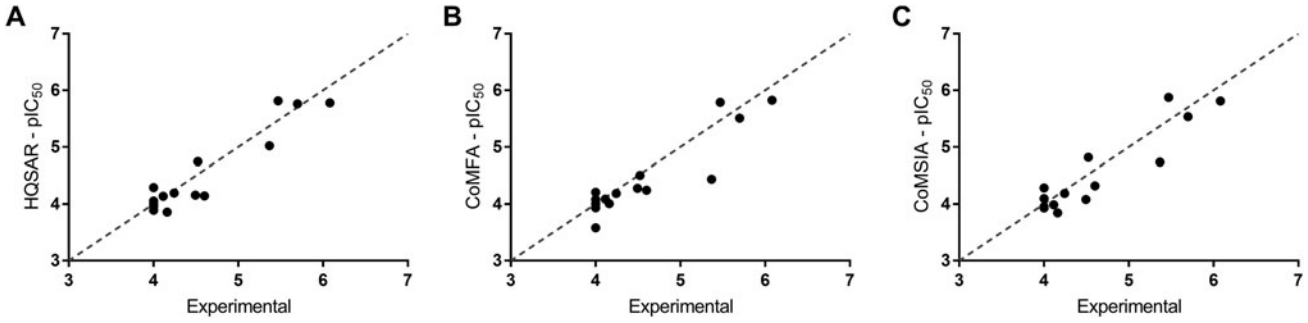


Figure 3. Correlation between experimentally determined and predicted pIC_{50} values for test set compounds, for individual techniques: HQSAR (A), CoMFA (B) and CoMSIA (C). The dashed line represents correlation equals one.

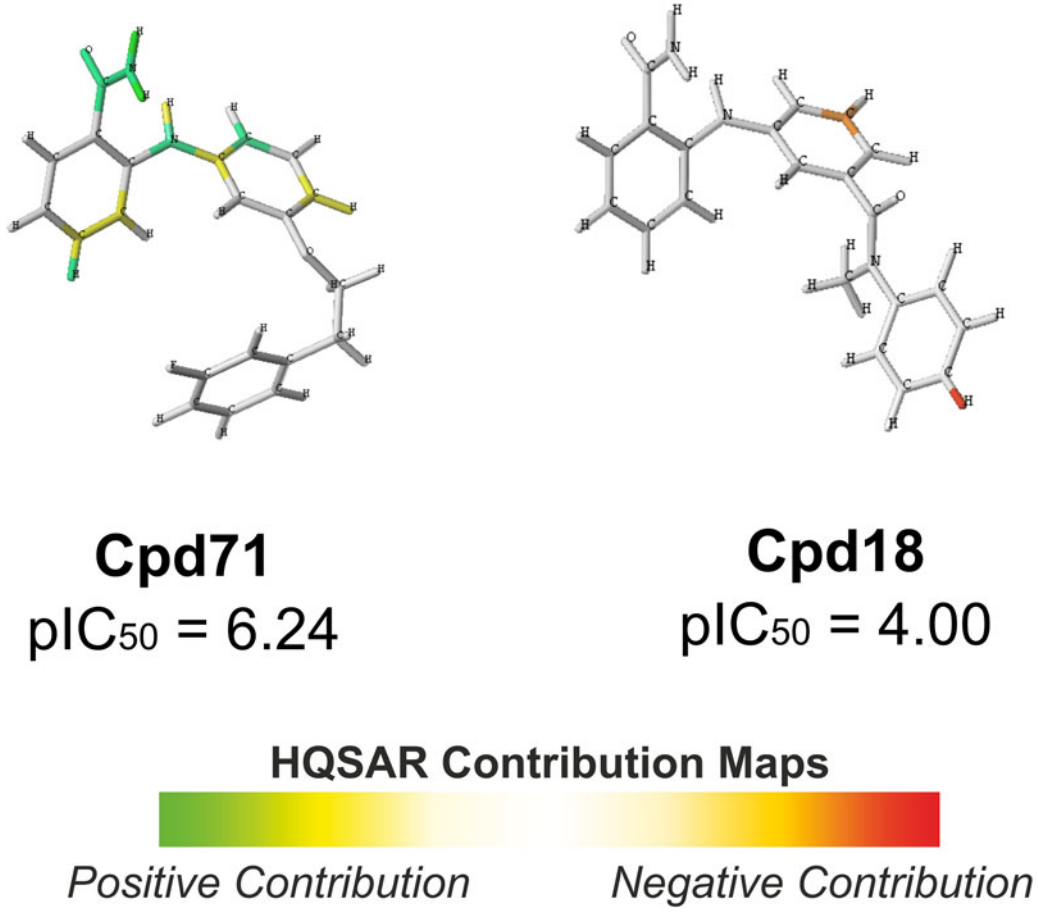


Figure 4. HQSAR contribution maps for the compounds 71, 18 and 9. Green, blue–green and yellow color atoms represent fragments with a positive contribution towards biological activity, while orange and red represent negative contribution (scale provided bellow structures).

for the compound structures) were essential to improve SIRT2 inhibition, with higher biological activity values compared to the other compounds. CoMFA suggests that charged moieties would be unfavorable for biological activity but could not confirm which would be the best substituent in this position. Remarkably, halogen substitution at the *meta* position of ring 3 leads to the most potent compound, specifically with fluorine (**71**, IC_{50} = 0.57 μM), compared to other positions (compounds **70** and **72**, with *ortho* and *para*, respectively), and with chlorine counterparts (compounds

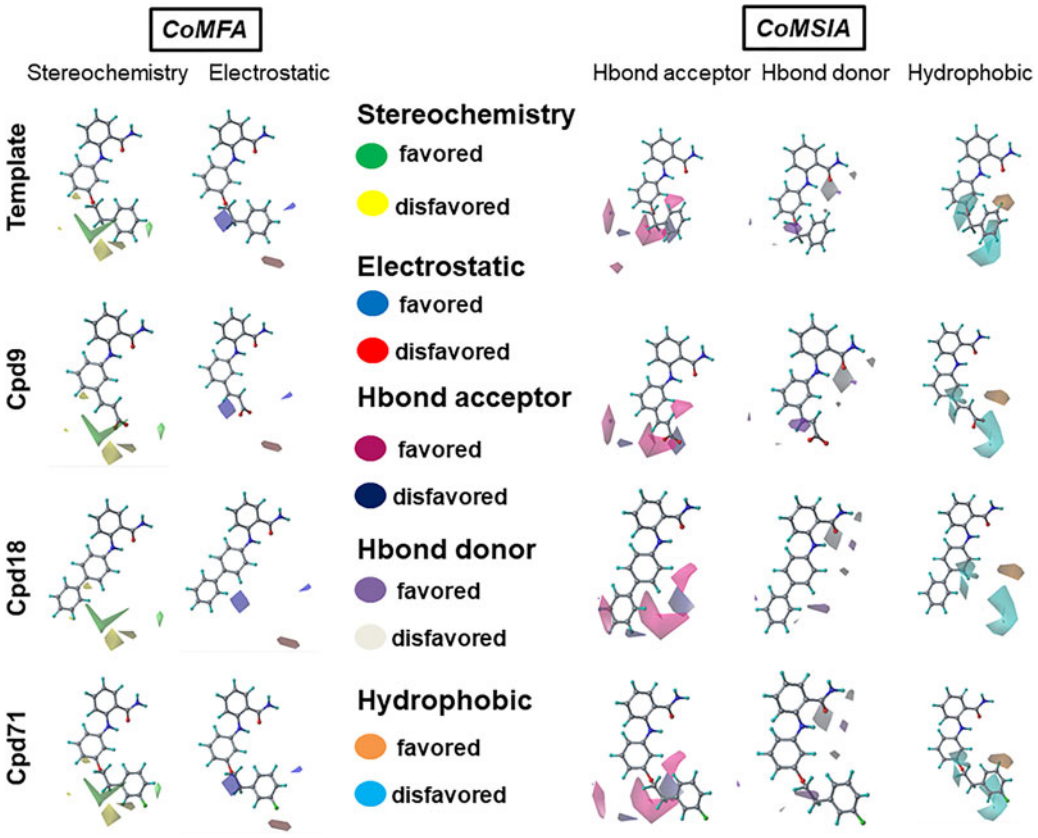


Figure 5. CoMFA contour maps, for steric and electrostatic terms, and CoMSIA contribution maps, highlighting the acceptor, donor and hydrophobic contributions. Contour maps showed around the compounds 71, 18 and 9. Green contours represent regions where bulky groups increase biological activity while yellow contours indicate areas where bulky groups decrease biological activity. Favorable electrostatic contributions represented in blue, while unfavorable contributions to the biological activity represented in red. Favorable acceptor contributions are highlighted in pink, while unfavorable in grey. Similarly, favorable donor contributions are highlighted in purple, while unfavorable in grey and favorable hydrophobic contributions are orange, while unfavorable are highlighted in cyan. Contour level: all favorable contours were set to 80 with their unfavorable counterparts set to 20.

73–75) there is an increase in activity; this further highlights the importance of the position of substitution for the activity of the molecule.

Docking of the anilinobenzamide within the SIRT2 active site

CoMFA models provided the relevance of hydrophobic substituents in the third ring. However, the precise influence of ring position could not be fully discussed by the ligand-based model; complementary docking calculations were employed to understand how those chemical features would interact within the protein structure. In this sense, molecular docking of relevant compounds (Figure 6A–C) highlighted the importance of Gln167, Asp95 and Val233 residues that, together with water molecules, are responsible for stabilization of polar portions of the compound; other studies corroborate this suggestion of network formation between water and residues for stabilization of the ligand (Lie, Thomsen, Pedersen, Schi⊘tt, & Christensen, 2011). Water molecules bridge the interaction of the compound's amide moiety with residues within the active site; A1112 interacts with the compound's nitrogen atom and main-chain oxygen atoms from His187 and Val233; A1129 interacts with the main chain of Gln167 and the charged side-chain of Asp95; and, finally, A1119 interacts with both oxygen atoms from A1129, the ligand's amide and the nitrogen from the main chain of Phe96. A comprehensive surface of hydrophobic and aromatic residues (Phe190, Phe131, Phe119, Phe235, Phe238 and Tyr139) creates an environment that favors hydrophobic and π-π type interactions, which fits the higher potency of ligands with higher apolar surfaces (Figure 6D). These characteristics can be employed to enrich compound libraries with representative testing candidates, applicable in an eventual virtual screening protocol or guide groups involved in chemical synthesis. Other related characteristics can be found in the Supporting Information (Figure 4).

Recent studies evolved a new selective inhibitor of SIRT2, namely KPM-2, with anticancer activity in the low micromolar band (Mellini et al., 2017; Rotili et al., 2012). KPM-2 can be conjugated by a covalent bond with NAD^+ within the active site of SIRT2, thus occupying the hydrophobic pocket (Phe96, Phe131, Phe143 and Phe190); such a portion is very structurally like the SIRT2 linker of Suzuki et al. (2012), where KPM-2 competes for the NAD^+ binding site. The KPM-2 double amide moiety has an additional hydrogen interaction with the Gly236 and Glu237 backbone, as well as the Gln267 side chain (Suzuki et al., 2012), as previously described (Schiedel et al., 2016). The free amide present in most of the

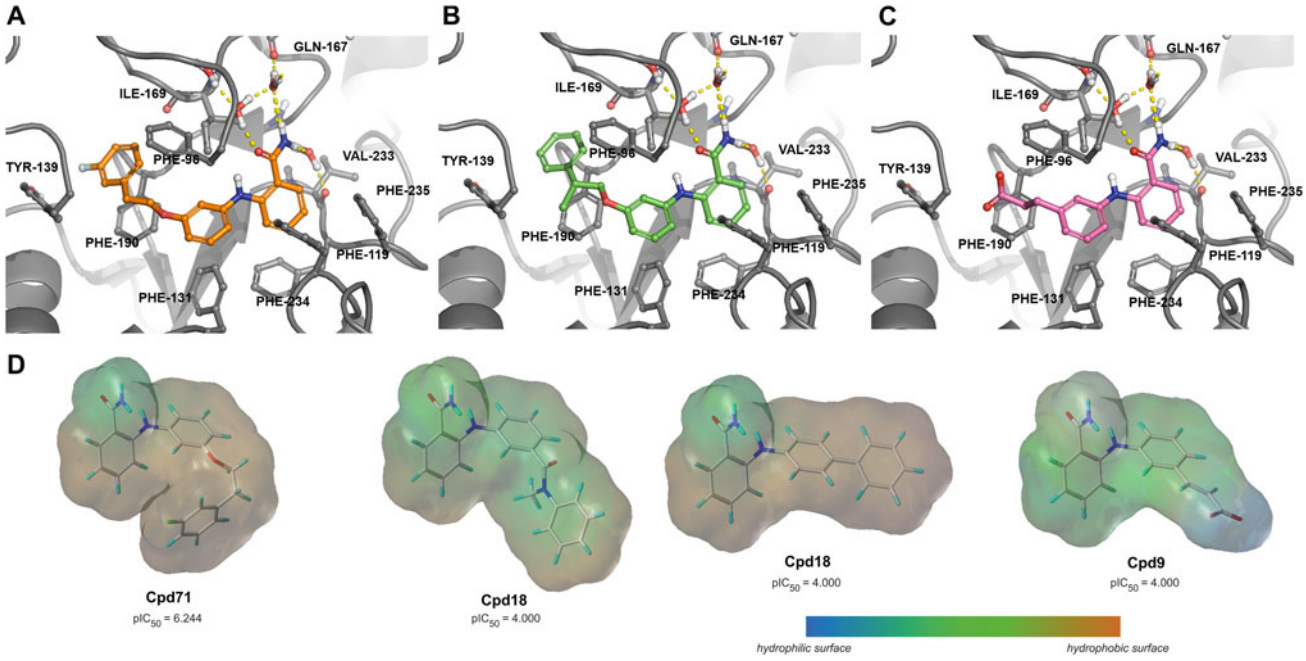


Figure 6. Molecular docking of representative compounds in the same pocket highlighting common interacting residues compounds **71** (A, orange), compound **18** (B, green) and compound **9** (C, pink). (D) Hydrophobicity map of representative dataset molecules.

compounds of the series can stabilize the network of hydrogen bond interactions that keep water molecules in position. This water network mediates the interaction with Gln267 in our model.

Additionally, the SIRT2 structure, co-crystallized with a myristoyl-H3K9 peptide (Protein Data Bank ID: 4X4O), shows that the fatty acid chain explores the hydrophobic groove composed of Phe96, Ile169, Phe119, Phe190 and Leu134 (Teng et al., 2015). Interestingly, new SIRT2 inhibitors, such as SirReal2 and AGK-2, can induce conformational changes in the protein that increase accessibility to this pocket, improving selectivity towards SIRT2 (Rumpf et al., 2015). The anilinobenzamide structural resemblance to those specific inhibitors suggests that they could also address SIRT2's hydrophobic pocket.

The best-generated HQSAR, CoMFA and CoMSIA models for predicting the SIRT2 enzyme were robust ($Q^2_{LOO} > 0.80$) and predictive ($r^2_{pred} > 0.70$). Physicochemical interpretation of the best models notes that insertion of a third ring and substitutions at *ortho*, *meta* and *para* positions contribute to the inhibitory activity of the compounds. Electrostatic maps of CoMFA and CoMSIA indicate possible interactions in these positions, with the most favorable being polar interactions with halogens F, Cl and Br in the *meta* position. Most active compounds have a clear separation between the hydrophilic region and the anilinobenzamide core, which can be beneficially extended with large hydrophobic groups (as discussed in the literature) to exploit a large specific pocket. As a future perspective, we hope that these QSAR models can provide insights into the design of new sirtuin-2-specific inhibitors.

Acknowledgments

The authors thank OpenEye Scientific Software for ROCS, OMEGA and QUACPAC academic licenses.

Author contribution

GMF, TK and VGM contributed to the experiment design. GMF, JGM, TK and VGM contributed to dataset organization. GMF and VGM contributed to model inference and validation. TK, GHT, AG, FSE and VGM contributed to discussion and writing of the paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the CNPq [grant number 312009/2014-3], [grant number 456984/2014-3], [grant number 310232/2017-1], [grant number 436791/2018-8]; FAPESP [grant number 2017/25543-8]; GlaxoSmithKline; and CAPES.

References

Alam, S., & Khan, F. (2017). 3D-QSAR studies on Maslinic acid analogs for anticancer activity against breast cancer cell line MCF-7. *Scientific Reports*, *7*(1), 6019. doi:10.1038/s41598-017-06131-0

Andreoli, F., & Del Rio, A. (2015). Computer-aided molecular design of compounds targeting histone modifying enzymes. *Computational and Structural Biotechnology Journal*, *13*, 358–365. doi:10.1016/j.csbj.2015.04.007

Bobrowska, A., Donmez, G., Weiss, A., Guarente, L., & Bates, G. (2012). SIRT2 ablation has no effect on tubulin acetylation in brain, cholesterol biosynthesis or the progression of Huntington's disease phenotypes in vivo. *PLoS One*, *7*(4), e34805. doi:10.1371/journal.pone.0034805

Braga, R. C., Alves, V. M., Silva, M. F. B., Muratov, E., Fourches, D., Tropsha, A., & Andrade, C. H. (2014). Tuning HERG out: antitarget QSAR models for drug development. *Current Topics in Medicinal Chemistry*, *14*(11), 1399–1415.

Chen, G., Zhou, D., Li, X.-Z., Jiang, Z., Tan, C., Wei, X.-Y., … Li, N. (2017). A natural chalcone induces apoptosis in lung cancer cells: 3D-QSAR,

docking and an in vivo/vitro assay. *Scientific Reports*, *7*(1), 10729. doi:10.1038/s41598-017-11369-9

Chirico, N., & Gramatica, P. (2011). Real external predictivity of QSAR models: how to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. *Journal of Chemical Information and Modeling*, *51*(9), 2320–2335. doi:10.1021/ci200211n

Choudhary, C., Weinert, B. T., Nishida, Y., Verdin, E., & Mann, M. (2014). The growing landscape of lysine acetylation links metabolism and cell signalling. *Nature Reviews Molecular Cell Biology*, *15*(8), 536–550. doi:10.1038/nrm3841

Chuang, Y.-C., Chang, C.-H., Lin, J.-T., & Yang, C.-N. (2015). Molecular modelling studies of sirtuin 2 inhibitors using three-dimensional structure–activity relationship analysis and molecular dynamics simulations. *Molecular Biosystems*, *11*(3), 723–733. doi:10.1039/C4MB00620H

Delano, W. L. (2004). *PyMOL reference guide*. San Carlos, California, USA: DeLano Scientific LLC.

Feinberg, A. P., & Vogelstein, B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature*, *301*(5895), 89–92. doi:10.1038/301089a0

Feldman, J. L., Baeza, J., & Denu, J. M. (2013). Activation of the protein deacetylase SIRT6 by long-chain fatty acids and widespread deacylation by mammalian sirtuins. *Journal of Biological Chemistry*, *288*(43), 31350–31356. doi:10.1074/jbc.C113.511261

Gaudio, A. C., & Zandonade, E. (2001). Proposição, validação e análise dos modelos que correlacionam estrutura química e atividade biológica. *Química Nova*, *24*(5), 658–671. doi:10.1590/S0100-40422001000500013

Gramatica, P. (2007). Principles of QSAR models validation: internal and external. *QSAR & Combinatorial Science*, *26*(5), 694–701. doi:10.1002/qsar.200610151

Halgren, T. A. (1996). Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, *17*(5–6), 490–519. doi:10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P

He, B., Hu, J., Zhang, X., & Lin, H. (2014). Thiomyristoyl peptides as cell-permeable Sirt6 inhibitors. Organic & Biomolecular Chemistry, *12*(38), 7498–7502. doi:10.1039/C4OB00860J

Infometrix. (1985). *Pirouette multivariate data analysis software*. Bothell, Washington: EUA.

Jiang, W., Wang, S., Xiao, M., Lin, Y., Zhou, L., Lei, Q., … Zhao, S. (2011). Acetylation regulates gluconeogenesis by promoting PEPCK1 degradation via recruiting the UBR5 ubiquitin ligase. *Molecular Cell*, *43*(1), 33–44. doi:10.1016/j.molcel.2011.04.028

Jing, E., Gesta, S., & Kahn, C. R. (2007). SIRT2 regulates adipocyte differentiation through FoxO1 acetylation/deacetylation. *Cell Metabolism*, *6*(2), 105–114. doi:10.1016/j.cmet.2007.07.003

Jones, G., Willett, P., Glen, R. C., Leach, A. R., & Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking 1 1Edited by F. E. Cohen. *Journal of Molecular Biology*, *267*(3), 727–748. doi:10.1006/jmbi.1996.0897

Kim, H.-S., Vassilopoulos, A., Wang, R.-H., Lahusen, T., Xiao, Z., Xu, X., … Deng, C.-X. (2011). SIRT2 maintains genome integrity and suppresses tumorigenesis through regulating APC/C activity. *Cancer Cell*, *20*(4), 487–499. doi:10.1016/j.ccr.2011.09.004

Kowalski, B. R. (Ed.). (1977). Chemometrics: Theory and application, vol. 52. Washington, DC: American Chemical Society. doi:10.1021/bk-1977-0052

Kunal, R., Supratik, K., & Das, R. N. (2015). Statistical methods in QSAR/QSPR. In E. Hawkins (Ed.), *A primer on QSAR/QSPR modeling – fundamental concepts* (pp. 37–59). New York: Springer International Publishing. doi:10.1007/978-3-319-17281-1

Lara, E., Mai, A., Calvanese, V., Altucci, L., Lopez-Nieva, P., Martinez-Chantar, M. L., … Fraga, M. F. (2009). Salermide, a Sirtuin inhibitor with a strong cancer-specific proapoptotic effect. *Oncogene*, *28*(6), 781–791. doi:10.1038/onc.2008.436

Lie, M. A., Thomsen, R., Pedersen, C. N. S., Schiøtt, B., & Christensen, M. H. (2011). Molecular docking with ligand attached water molecules. *Journal of Chemical Information and Modeling*, *51*(4), 909–917. doi:10.1021/ci100510m

Lin, R., Tao, R., Gao, X., Li, T., Zhou, X., Guan, K.-L., … Lei, Q.-Y. (2013). Acetylation stabilizes ATP-citrate lyase to promote lipid biosynthesis and tumor growth. *Molecular Cell*, *51*(4), 506–518. doi:10.1016/j.molcel.2013.07.002

Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design*, *27*(3), 221–234. doi:10.1007/s10822-013-9644-8

Mai, A., Massa, S., Pezzi, R., Valente, S., Loidl, P., & Brosch, G. (2005). Synthesis and biological evaluation of 2-, 3-, and 4-acylaminocinnamyl-N-hydroxyamides as novel synthetic HDAC inhibitors. *Medicinal Chemistry (Shariqah (United Arab Emirates))*, *1*(3), 245–254.

Mellini, P., Itoh, Y., Tsumoto, H., Li, Y., Suzuki, M., Tokuda, N., … Suzuki, T. (2017). Potent mechanism-based sirtuin-2-selective inhibition by an in situ-generated occupant of the substrate-binding site, "selectivity pocket" and NAD+-binding site. *Chemical Science*, *8*(9), 6400–6408. doi:10.1039/C7SC02738A

Menzies, K. J., Zhang, H., Katsyuba, E., & Auwerx, J. (2016). Protein acetylation in metabolism - metabolites and cofactors. *Nature Reviews Endocrinology*, *12*(1), 43–60. doi:10.1038/nrendo.2015.181

Napper, A. D., Hixon, J., McDonagh, T., Keavey, K., Pons, J. F., Barker, J. … , Curtis, R. (2005). Discovery of indoles as potent and selective inhibitors of the deacetylase SIRT1. *Journal of Medicinal Chemistry*, *48*(25), 8045–8054. doi:10.1021/JM050522V

Puigserver, P., Wu, Z., Park, C. W., Graves, R., Wright, M., & Spiegelman, B. M. (1998). A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell*, *92*(6), 829–839.(00)81410-5 doi:10.1016/S0092-8674(00)81410-5

Ramis, M. R., Esteban, S., Miralles, A., Tan, D. X., & Reiter, R. J. (2015). Caloric restriction, resveratrol and melatonin: Role of SIRT1 and implications for aging and related-diseases. *Mech Ageing Dev*, *146–148*, 28–41. doi:10.1016/j.mad.2015.03.008

Rodgers, J. T., Lerin, C., Haas, W., Gygi, S. P., Spiegelman, B. M., & Puigserver, P. (2005). Nutrient control of glucose homeostasis through a complex of PGC-1α and SIRT1. *Nature*, *434*(7029), 113–118. doi:10.1038/nature03354

Rodriquez, M., Aquino, M., Bruno, I., De Martino, G., Taddei, M., & Gomez-Paloma, L. (2006). Chemistry and biology of chromatin remodeling agents: state of art and future perspectives of HDAC inhibitors. *Current Medicinal Chemistry*, *13*(10), 1119–1139. doi:10.2174/092986706776360905

Rotili, D., Tarantino, D., Nebbioso, A., Paolini, C., Huidobro, C., Lara, E., … Mai, A. (2012). Discovery of salermide-related sirtuin inhibitors: binding mode studies and antiproliferative effects in cancer cells including cancer stem cells. *Journal of Medicinal Chemistry*, *55*(24), 10937–10947. doi:10.1021/jm3011614

Roy, K., Das, R. N., Ambure, P., & Aher, R. B. (2016). Be aware of error measures. Further studies on validation of predictive QSAR models. *Chemometrics and Intelligent Laboratory Systems*, *152*, 18–33. doi:10.1016/j.chemolab.2016.01.008

Rumpf, T., Schiedel, M., Karaman, B., Roessler, C., North, B. J., Lehotzky, A., … Jung, M. (2015). Selective Sirt2 inhibition by ligand-induced rearrangement of the active site. *Nature Communications*, *6*(1), 62–63. doi:10.1038/ncomms7263

Satoh, A., Vudikaria, L. Q., Kurano, N., & Miyachi, S. (2005). Evaluation of the sensitivity of marine microalgal strains to the heavy metals, Cu, As, Sb, Pb and Cd. *Environ Int*, *31*(5), 713–722. doi:10.1016/j.envint.2005.01.001

Schiedel, M., Rumpf, T., Karaman, B., Lehotzky, A., Gerhardt, S., Ovádi, J., … Jung, M. (2016). Structure-based development of an affinity probe for Sirtuin 2. *Angewandte Chemie International Edition* , *55*(6), 2252–2256. doi:10.1002/anie.201509843

Seel, M., Turner, D. B., & Willett, P. (1999). Effect of parameter variations on the effectiveness of HQSAR analyses. *Quantitative Structure-Activity Relationships*, *18*(3), 245–252.doi:10.1002/(SICI)1521-3838(199907)18:3<245::AID-QSAR245>3.0.CO;2-O

Serrano, L., Martínez-Redondo, P., Marazuela-Duque, A., Vazquez, B. N., Dooley, S. J., Voigt, P., … Vaquero, A. (2013). The tumor suppressor SirT2 regulates cell cycle progression and genome stability by modulating the mitotic deposition of H4K20 methylation. *Genes & Development*, *27*(6), 639–653. doi:10.1101/gad.211342.112

Siligardi, G., Hussain, R., Patching, S. G., & Phillips-Jones, M. K. (2014). Ligand-and drug-binding studies of membrane proteins revealed through circular dichroism spectroscopy. *Biochimica Biophysics Acta*, *1838*, 34–42. doi:10.1016/j.bbamem.2013.06.019

Suzuki, T., Imai, K., Imai, E., Iida, S., Ueda, R., Tsumoto, H., … Miyata, N. (2009). Design, synthesis, enzyme inhibition, and tumor cell growth inhibition of 2-anilinobenzamide derivatives as SIRT1 inhibitors. *Bioorganic & Medicinal Chemistry*, *17*(16), 5900–5905. doi:10.1016/j.bmc.2009.07.001

Suzuki, T., Khan, M. N. A., Sawada, H., Imai, E., Itoh, Y., Yamatsuta, K., … Miyata, N. (2012). Design, synthesis, and biological activity of a novel series of human Sirtuin-2-selective inhibitors. *Journal of Medicinal Chemistry*, *55*(12), 5760–5773. doi:10.1021/jm3002108

Teng, Y.-B., Jing, H., Aramsangtienchai, P., He, B., Khan, S., Hu, J., … Hao, Q. (2015). Efficient demyristoylase activity of SIRT2 revealed by kinetic and structural studies. *Scientific Reports*, *5*(1), 8529. doi:10.1038/srep08529

Tripos, Inc. (2011). *SYBYL ® expert molecular modeling environment*.

van de Ven, R. A. H., Santos, D., & Haigis, M. C. (2017). Mitochondrial Sirtuins and molecular mechanisms of aging. *Trends in Molecular Medicine*, *23*(4), 320–331. doi:10.1016/j.molmed.2017.02.005

Verdonk, M. L., Chessari, G., Cole, J. C., Hartshorn, M. J., Murray, C. W., Nissink, J. W., …., Taylor, R. (2005). Modeling water molecules in protein − ligand docking using GOLD. *Journal of Medicinal Chemistry*, *48*(20), 6504–6515. doi:10.1021/JM050543P

Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W., & Taylor, R. D. (2003). Improved protein-ligand docking using GOLD. *Proteins: Structure, Function, and Bioinformatics*, *52*(4), 609–623. doi:10.1002/prot.10465

Wilkes, J. G., Stoyanova-Slavova, I. B., & Buzatu, D. A. (2016). Alignment-independent technique for 3D QSAR analysis. *Journal of Computer-Aided Molecular Design*, *30*(4), 331–345. doi:10.1007/s10822-016-9909-0

Yang, M. H., Laurent, G., Bause, A. S., Spang, R., German, N., Haigis, M. C., & Haigis, K. M. (2013). HDAC6 and SIRT2 regulate the acetylation state and oncogenic activity of mutant K-RAS. *Molecular Cancer Research*, *11*(9), 1072–1077. doi:10.1158/1541-7786.MCR-13-0040-T

Yap, C. W. (2011). PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *Journal of Computational Chemistry*, *32*(7), 1466–1474. doi:10.1002/jcc.21707

Yoon, J. C., Puigserver, P., Chen, G., Donovan, J., Wu, Z., Rhee, J., … Spiegelman, B. M. (2001). Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. *Nature*, *413*(6852), 131–138. doi:10.1038/35093050

Zhao, M., Wang, L., Zheng, L., Zhang, M., Qiu, C., Zhang, Y., … Niu, B. (2017). 2D-QSAR and 3D-QSAR analyses for EGFR inhibitors. *BioMed Research International*, *2017*, 4649191. doi:10.1155/2017/4649191